ML application in ATLAS and CMS HH analyses

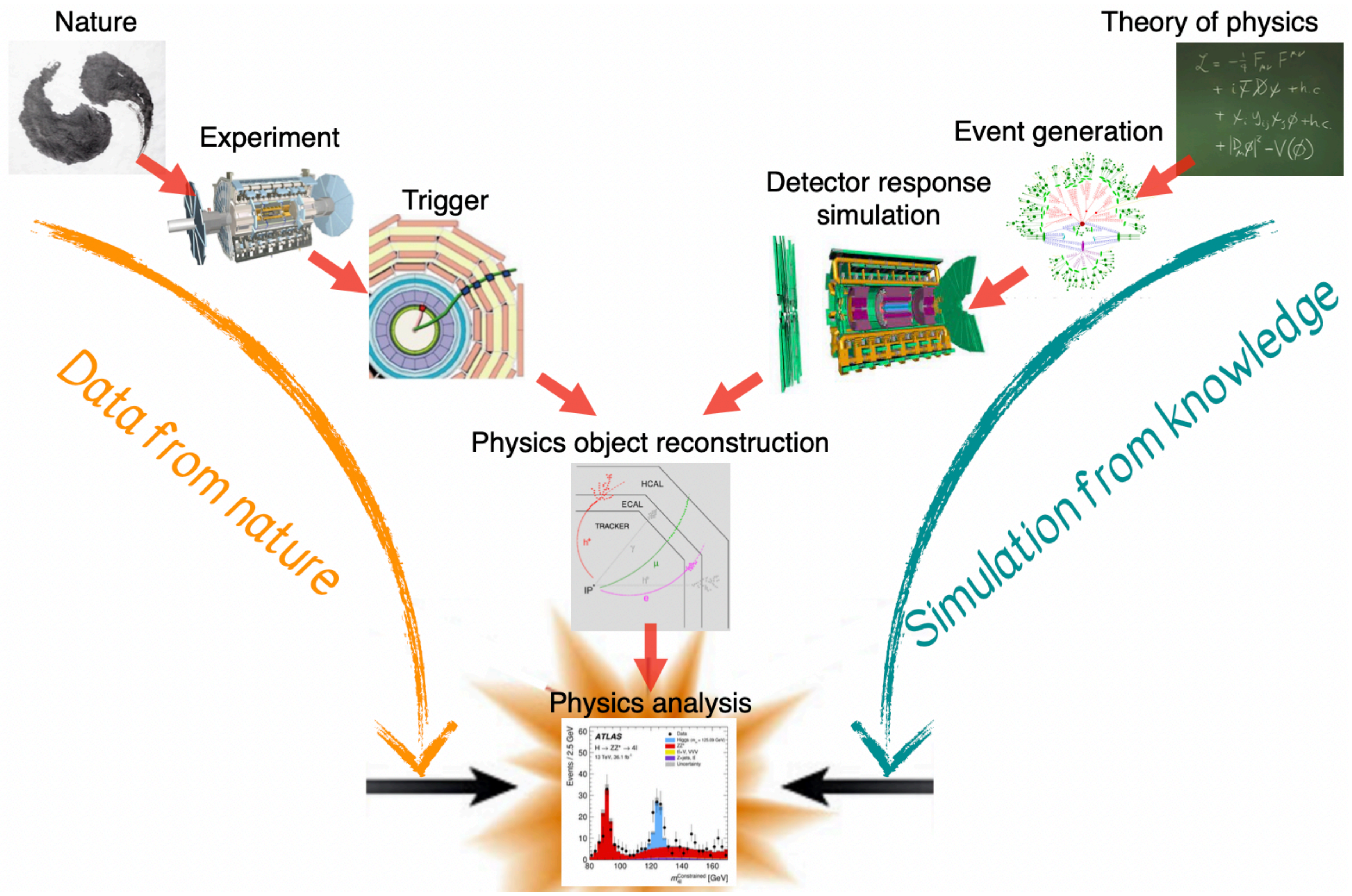
Chen Zhou (周辰)

Peking University (北京大学)

Higgs Potential Workshop
2024.12.21

Introduction

- One of the major objectives of the experimental programs at the LHC is the discovery of new physics
- Machine Learning: “application of artificial intelligence that provides systems the ability to automatically learn and improve from experience without being explicitly programmed”
  - It has become one of the most powerful techniques for High Energy Physics (HEP) data analysis
  - **It greatly enhances our ability of identifying signal from background: important for discovery of HH**



Back to Higgs discovery era

CMS H→γγ analysis (2012)

Select events with **two photons**

➔Train **Diphoton MVA** using signal and background MC

- Input variables: kinematics and (BDT-based) photon ID MVA of each photon, (BDT-based) vertex probability, etc.

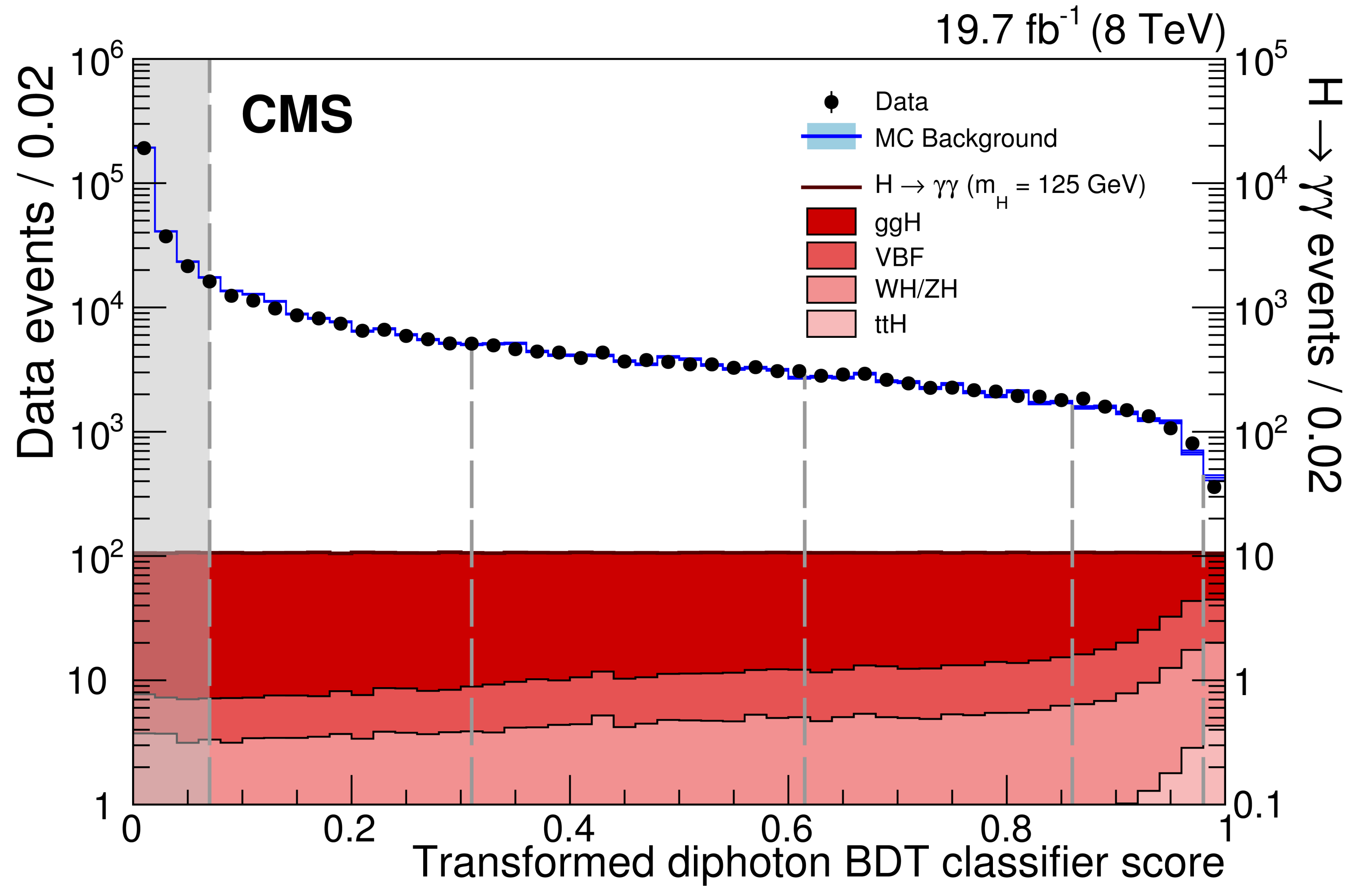
➔ Separate events to **categories** based on BDT score (which is to the first order independent of diphoton mass)

➔Fit **diphoton mass** over all categories

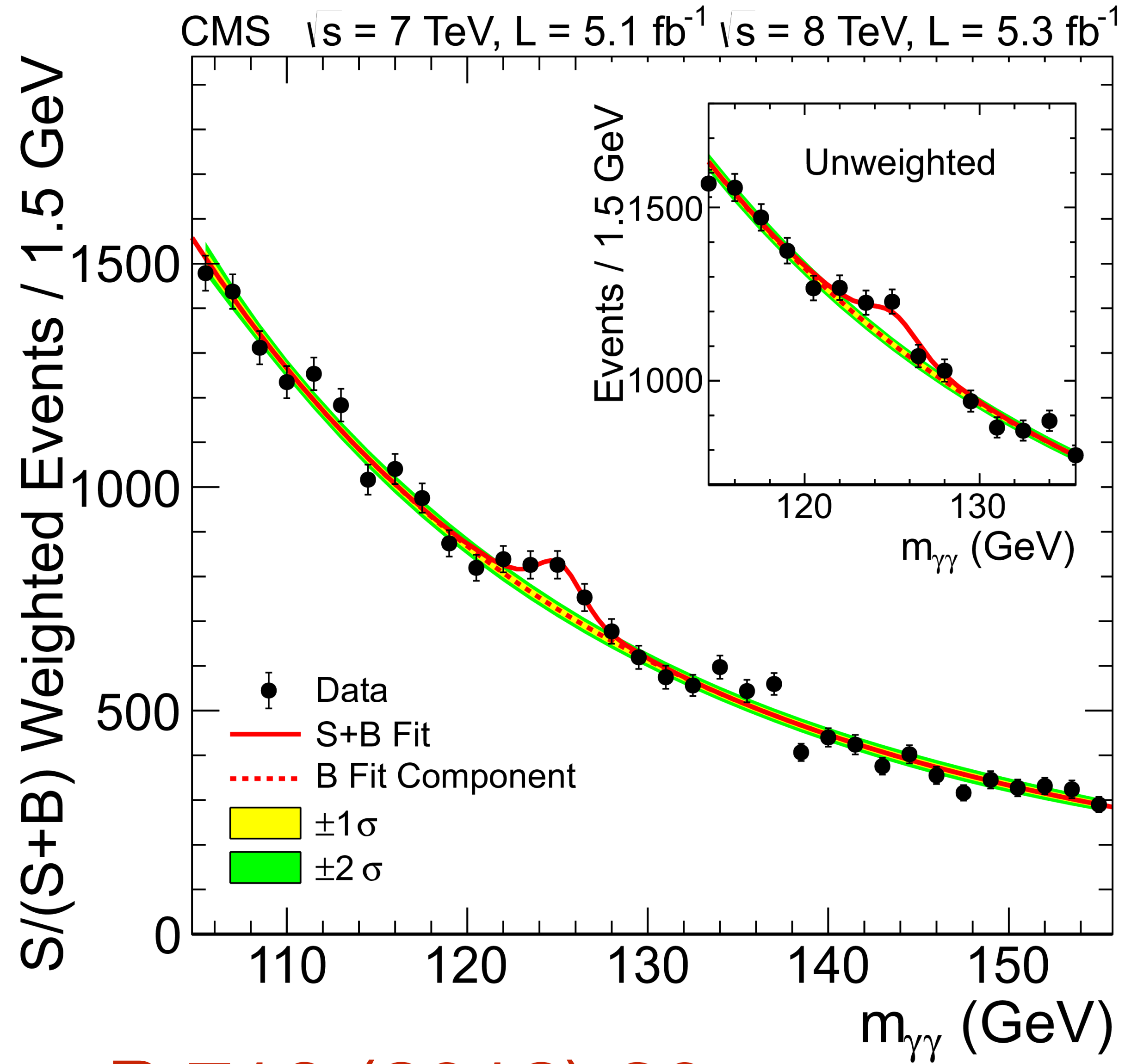
- Signature: a narrow resonance above a smooth background (QCD γγ production, etc.)

➔Measure signal strength, etc.

Better than cut-based analysis by 15%



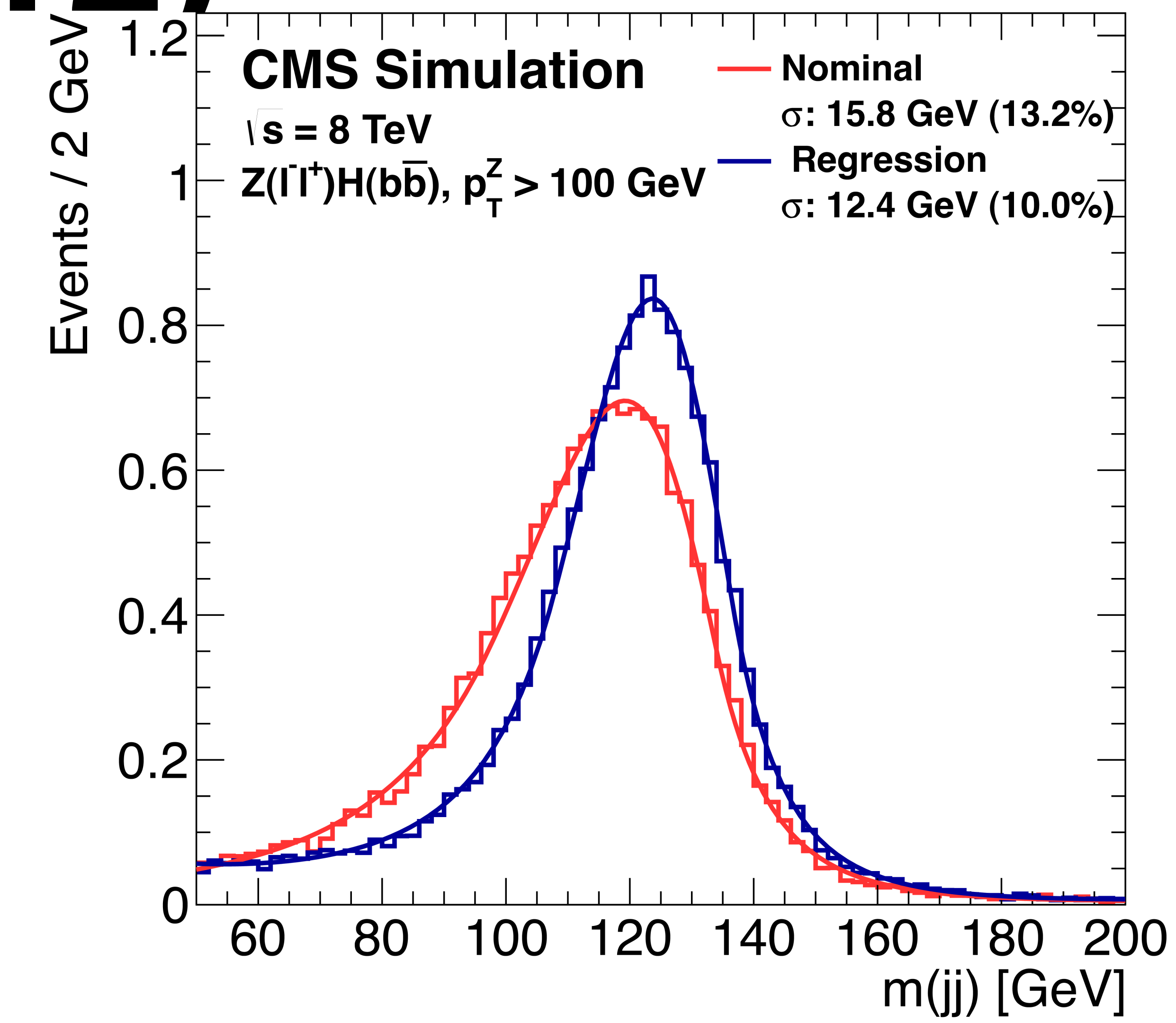
Eur. Phys. J. C 74 (2014) 3076



Phys. Lett. B 716 (2012) 30

CMS H→b$\bar{b}$ analysis (2012)

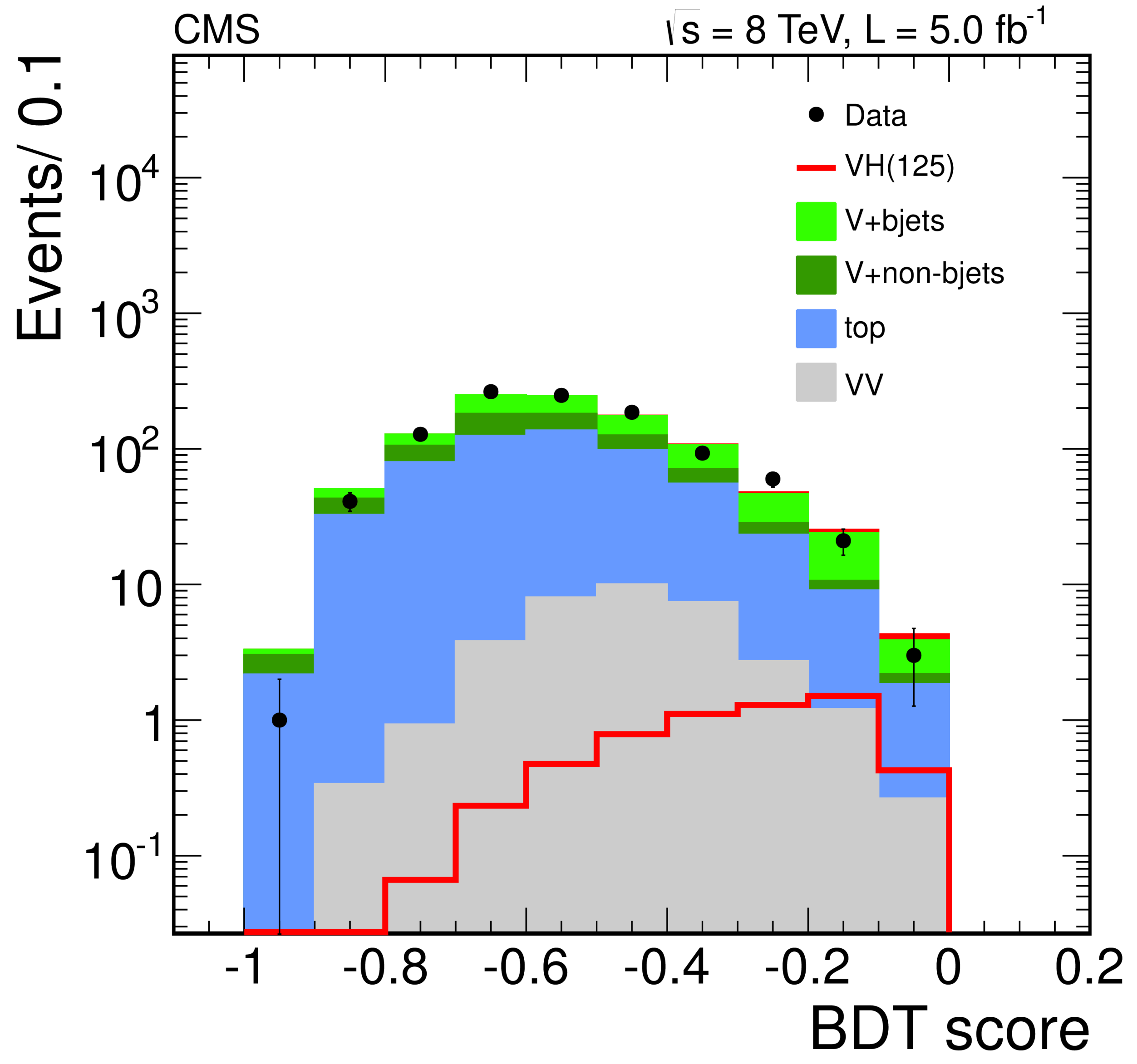
- **Large branch ratio (~58%)**
- Huge background, tackled by requiring associated particles and machine learning



Phys. Rev. D 89 (2014) 012003

VH→Vb$\bar{b}$

- Reconstruct Higgs as two small-radius b-tagged jets
- Tag leptonically decaying W/Z boson
- Main bkg: V+heavy flavor, t$\bar{t}$
- Train BDTs using kinematics of V and H candidates (e.g. $m_{b\bar{b}}$ reconstructed by regression)
- Fit the shape of the BDT output distribution

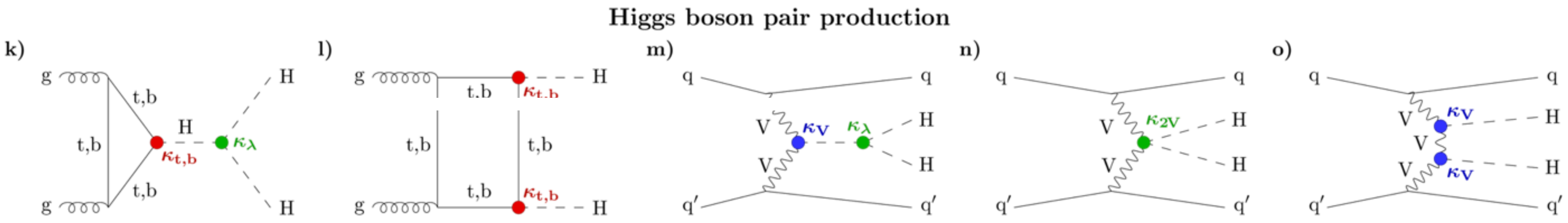


Phys. Lett. B 716 (2012) 30

AI-based event classification in Run-2 HH analyses

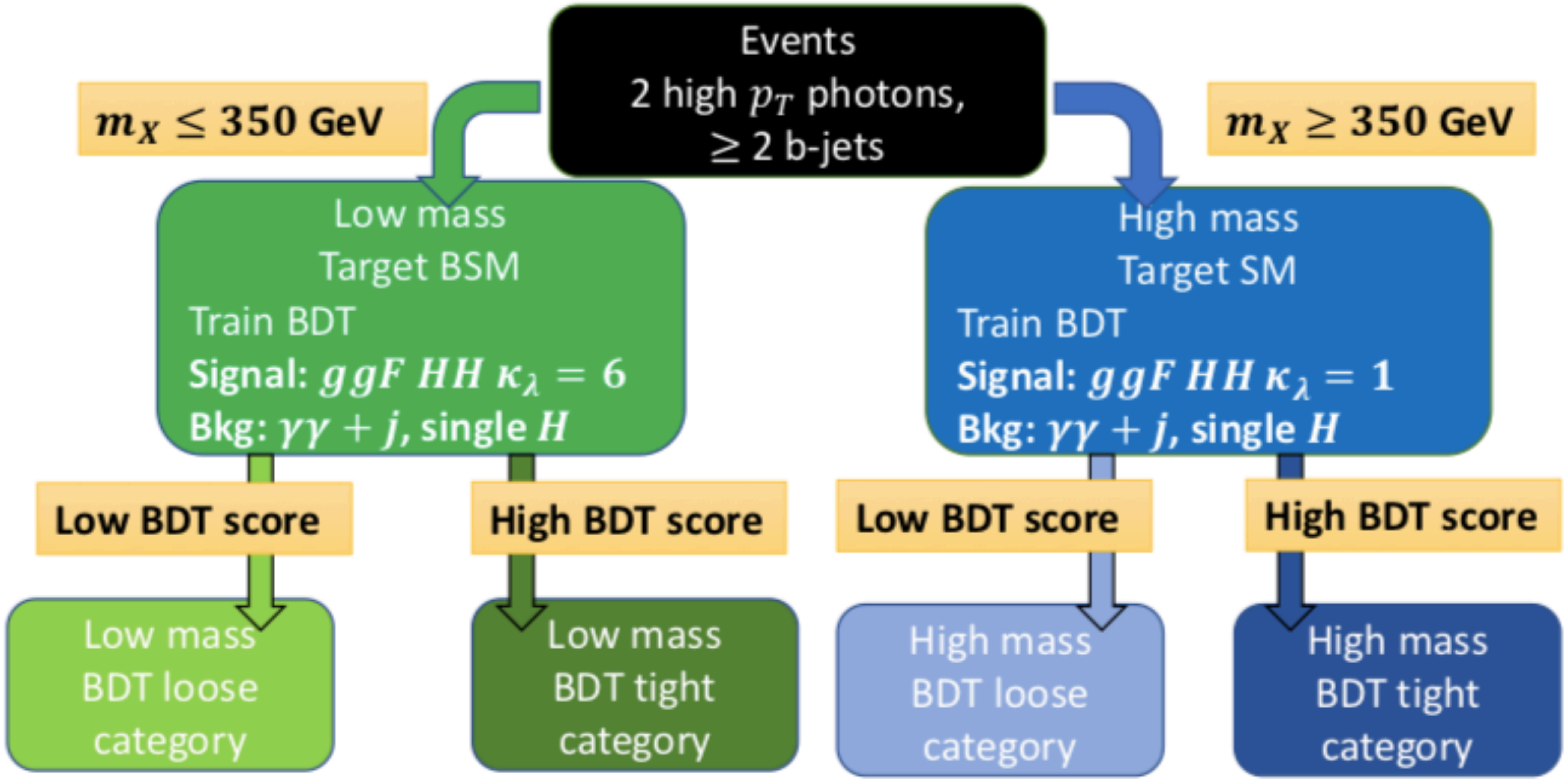
Higgs boson self-couplings

- **Higgs self-coupling is one of the deepest questions of SM and may provide a portal to new physics beyond it**
  - Vacuum stability, early universe evolvement, …
- **Double Higgs production is the way to directly probe Higgs self-couplings at the LHC**
  - Extremely low cross-section in the SM
  - Non-SM self-coupling strength can change cross-section and kinematics of double Higgs production



ATLAS HH→bbγγ analysis (2023)

JHEP 01 (2024) 066

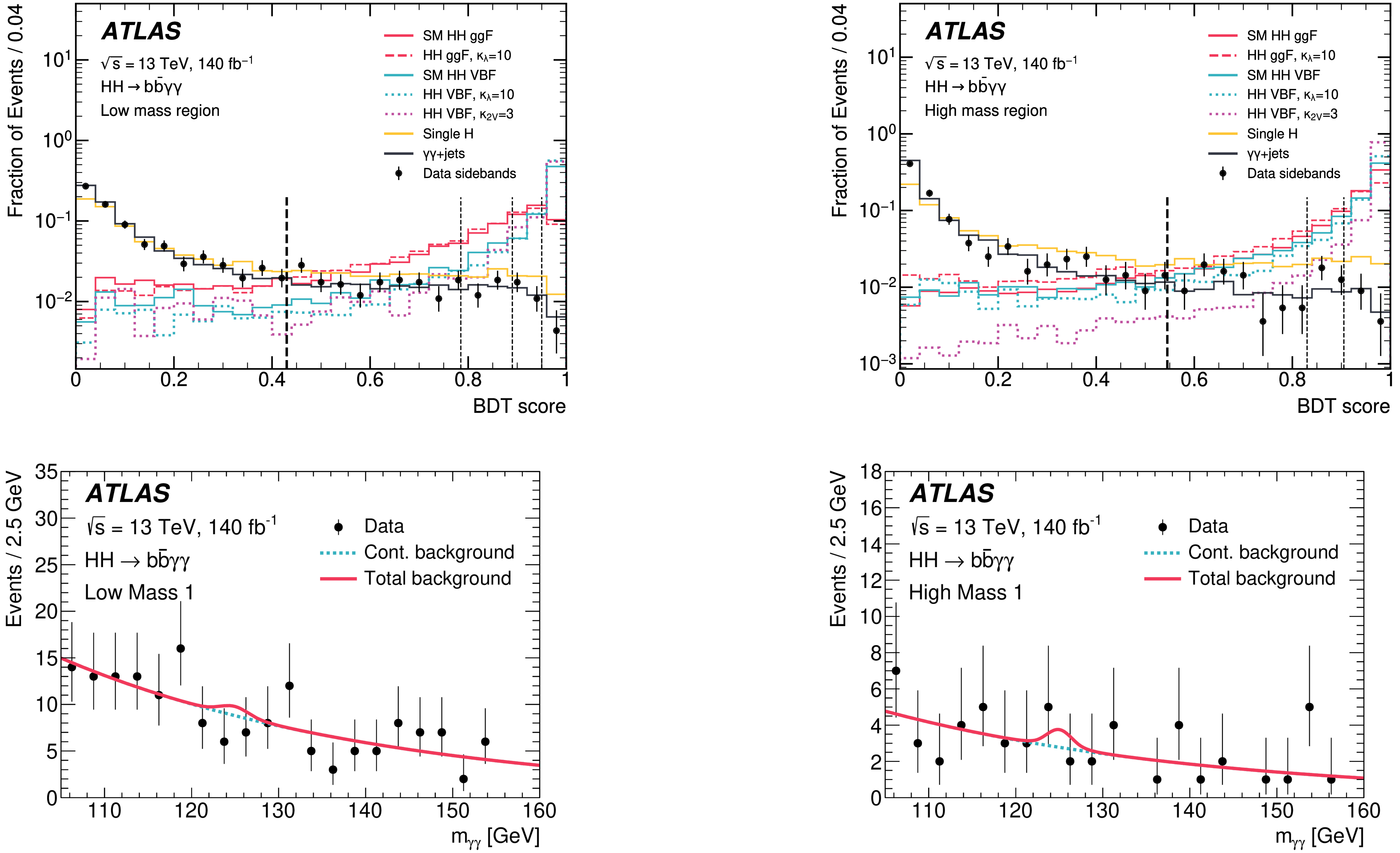


- BDT is trained with XGBoost to distinguish between HH and background
- Inputs include kinematic variables of photons & b-jets, as well as:
  - kinematic variables of VBF-jets which are identified by BDT-based tagger
  - event-level variables such as mass(bbγγ) and “topness”

$$\chi_{Wt} = \min\sqrt{\left(\frac{m_{j_1 j_2} - m_W}{m_W}\right)^2 + \left(\frac{m_{j_1 j_2 j_3} - m_t}{m_t}\right)^2},$$

ATLAS HH→bbγγ analysis (2023)

JHEP 01 (2024) 066

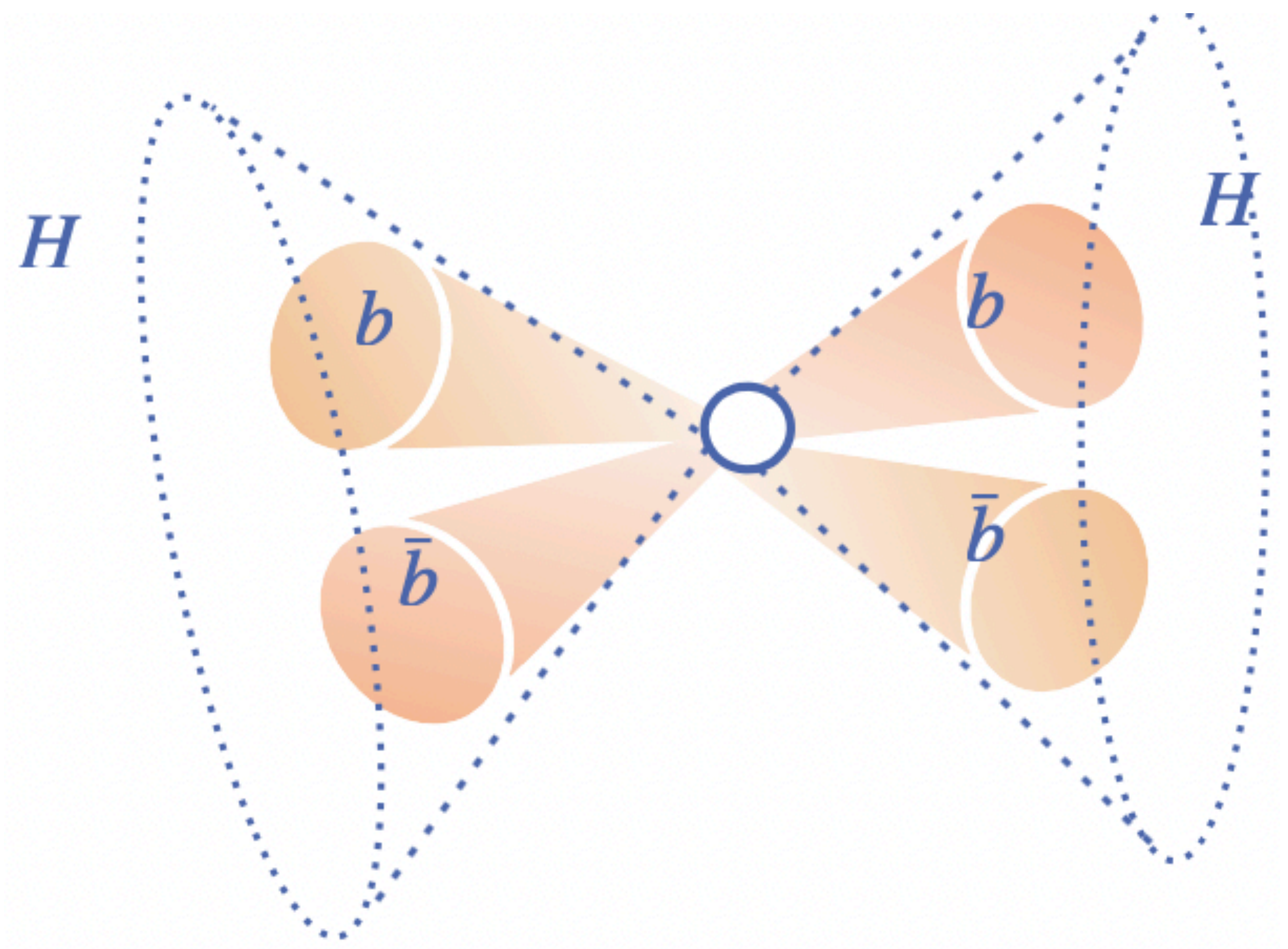


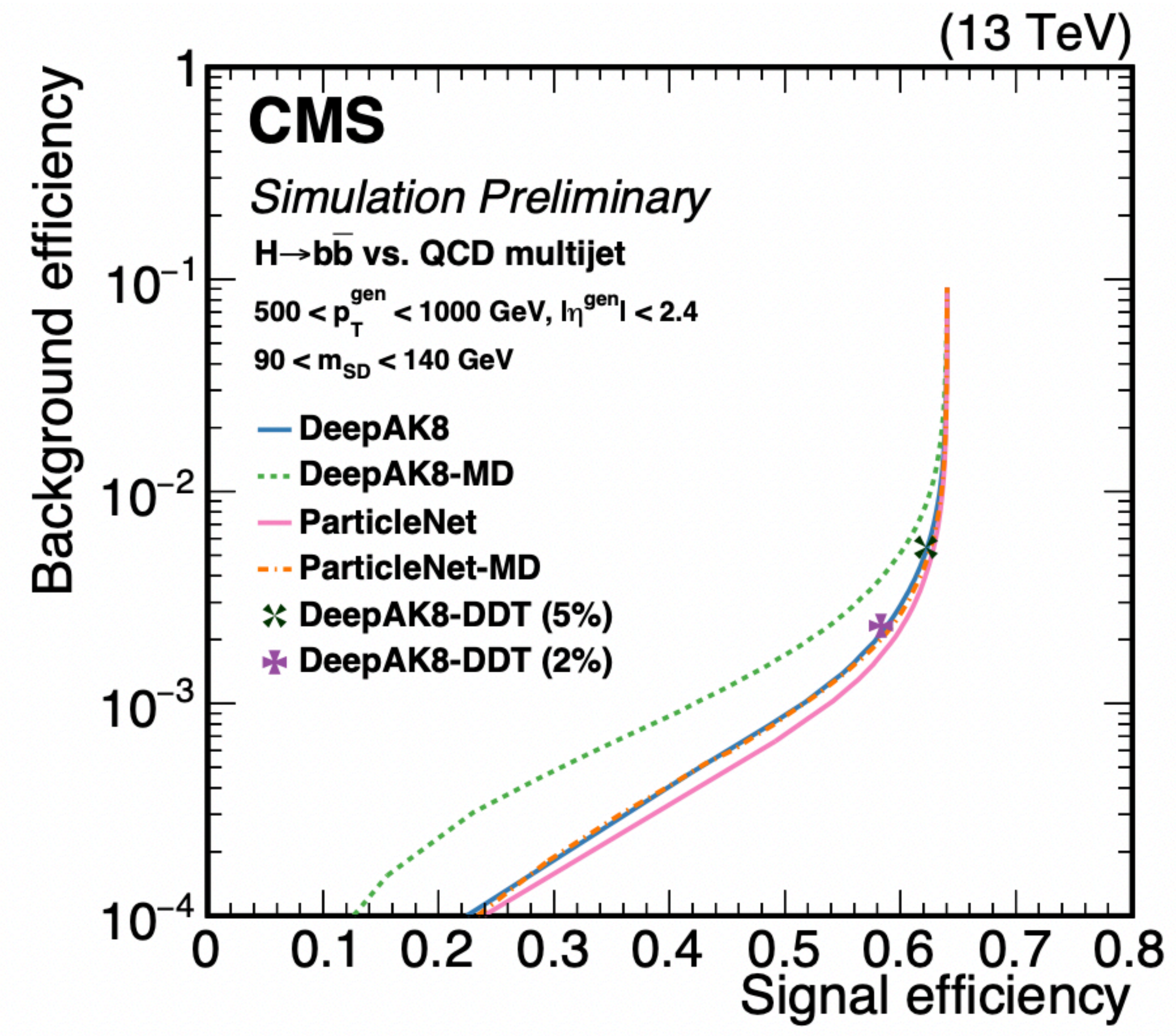
- Both HHH and HHVV couplings are optimized
- HHH coupling strength: $-1.4<\kappa_\lambda<6.9$; HHVV coupling strength: $-0.5<\kappa_{2V}<2.7$

Use of deep learning with low-level inputs

CMS non-resonant boosted HH→bbbb analysis (2022)

- Focus on phase space region where both Higgs bosons are highly Lorentz boosted
- Reconstruction and identification of b quark pair from Higgs decay is achieved with **ParticleNet, a graph neural network algorithm**
  - Using PF candidates and secondary vertices as inputs, yielding substantial gains over other approaches

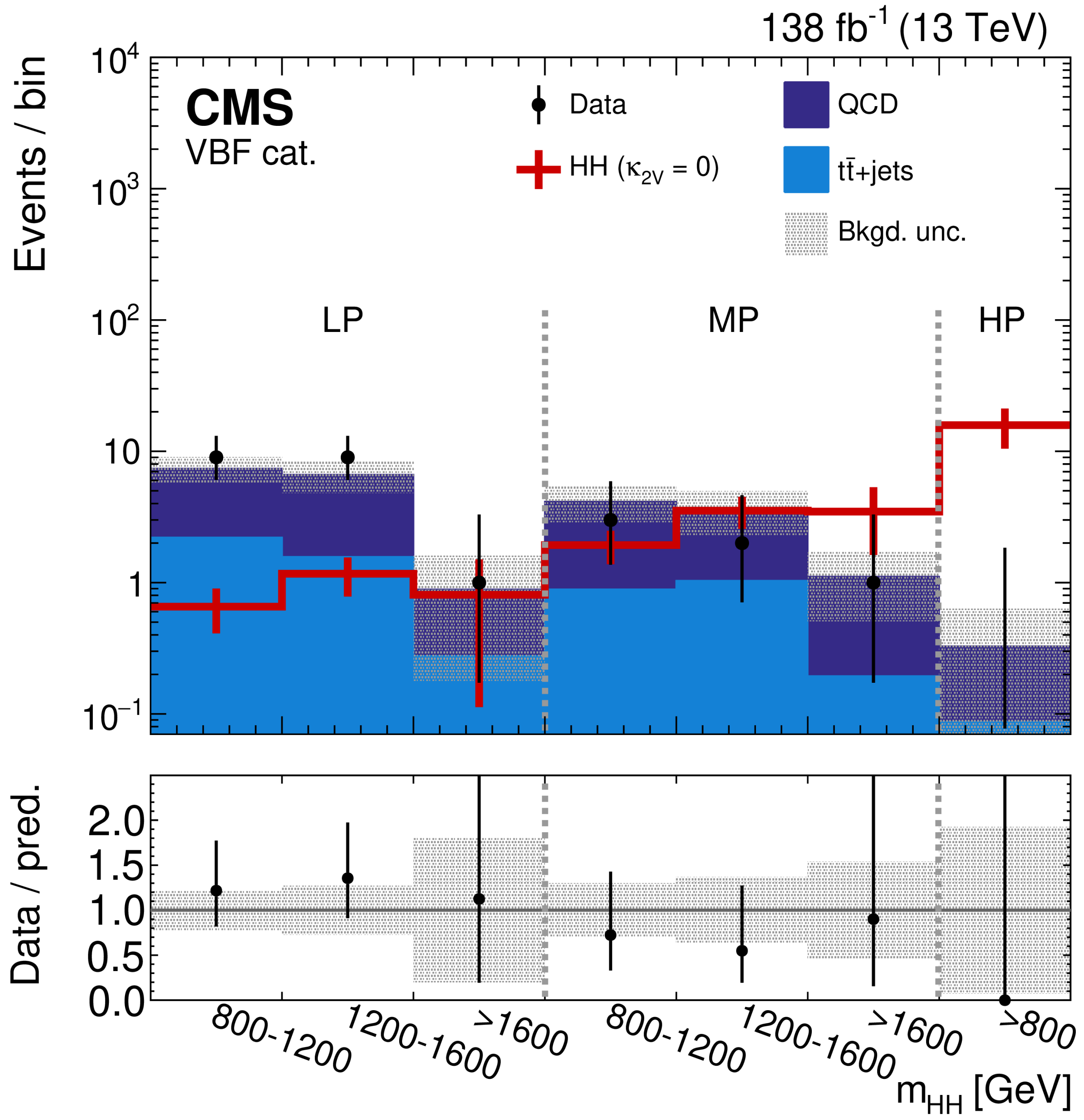


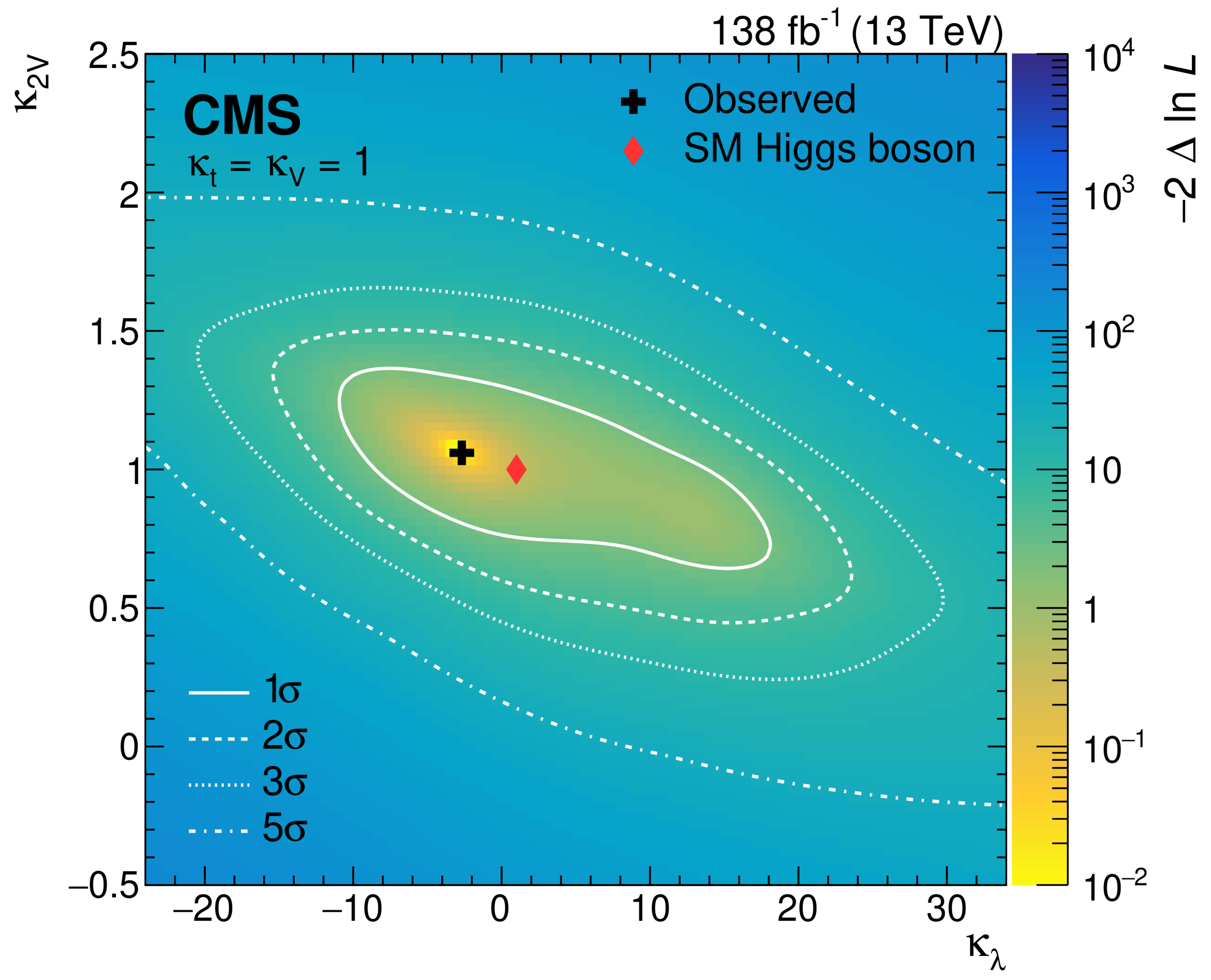


Phys. Rev. Lett. 131 (2023) 041803

CMS non-resonant boosted HH→bbbb analysis (2022)

- HH candidate mass is taken as final discriminant
- Constrains the H self-coupling strength and the quartic VVHH coupling strength κ_{2V}
  - **Excluding κ_{2V}=0 for the first time**, with a significance of 6.3σ





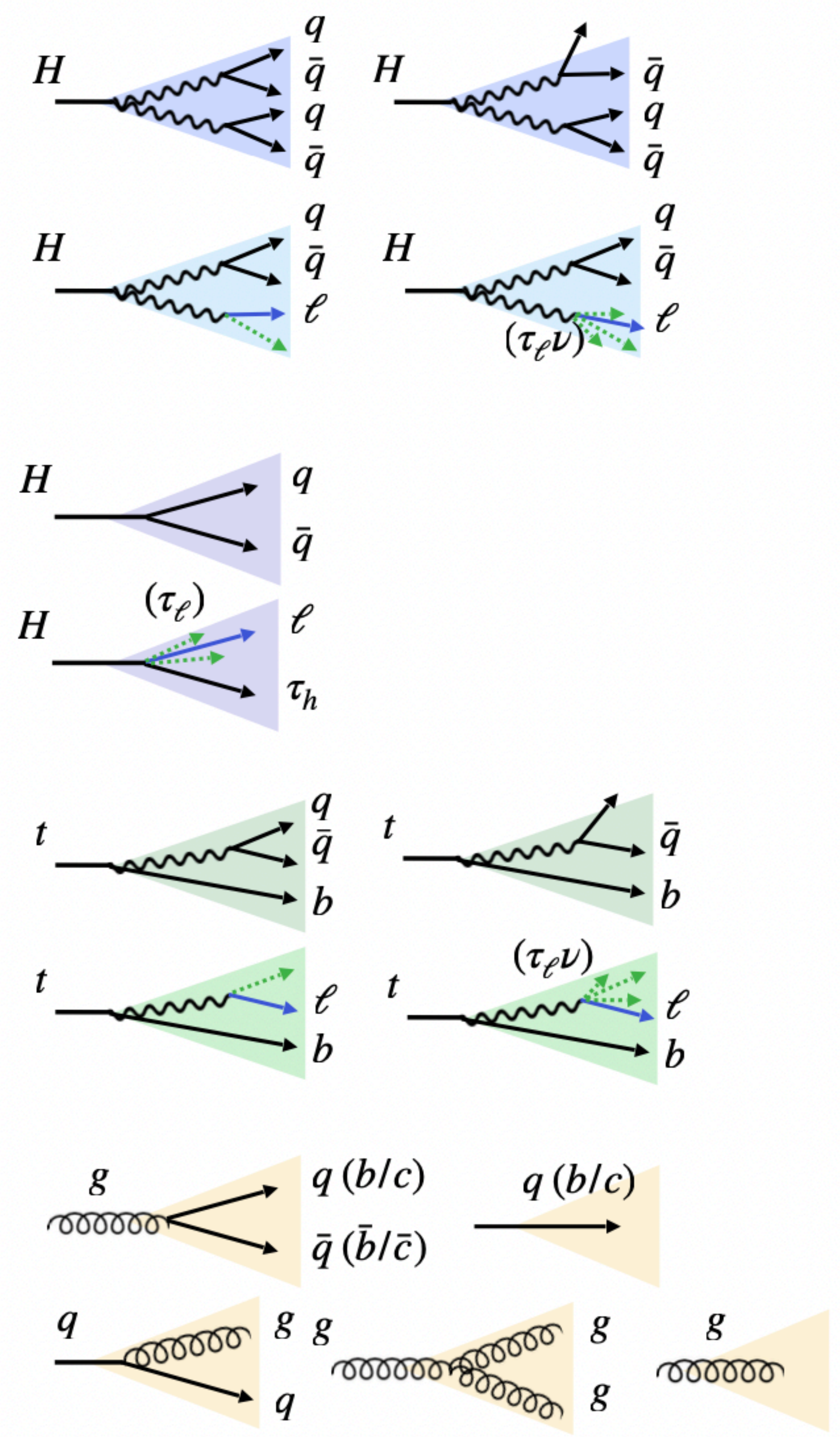
Phys. Rev. Lett. 131 (2023) 041803

CMS non-resonant boosted HH→bbVV analysis (2024)

- Extend to a large array of final states, including H→VV, all-hadronic, and semi-leptonic modes

- **Global Particle Transformer algorithm (GloParT)** uses learned “attention” to give more weight to certain particles in order to infer the origin of jets

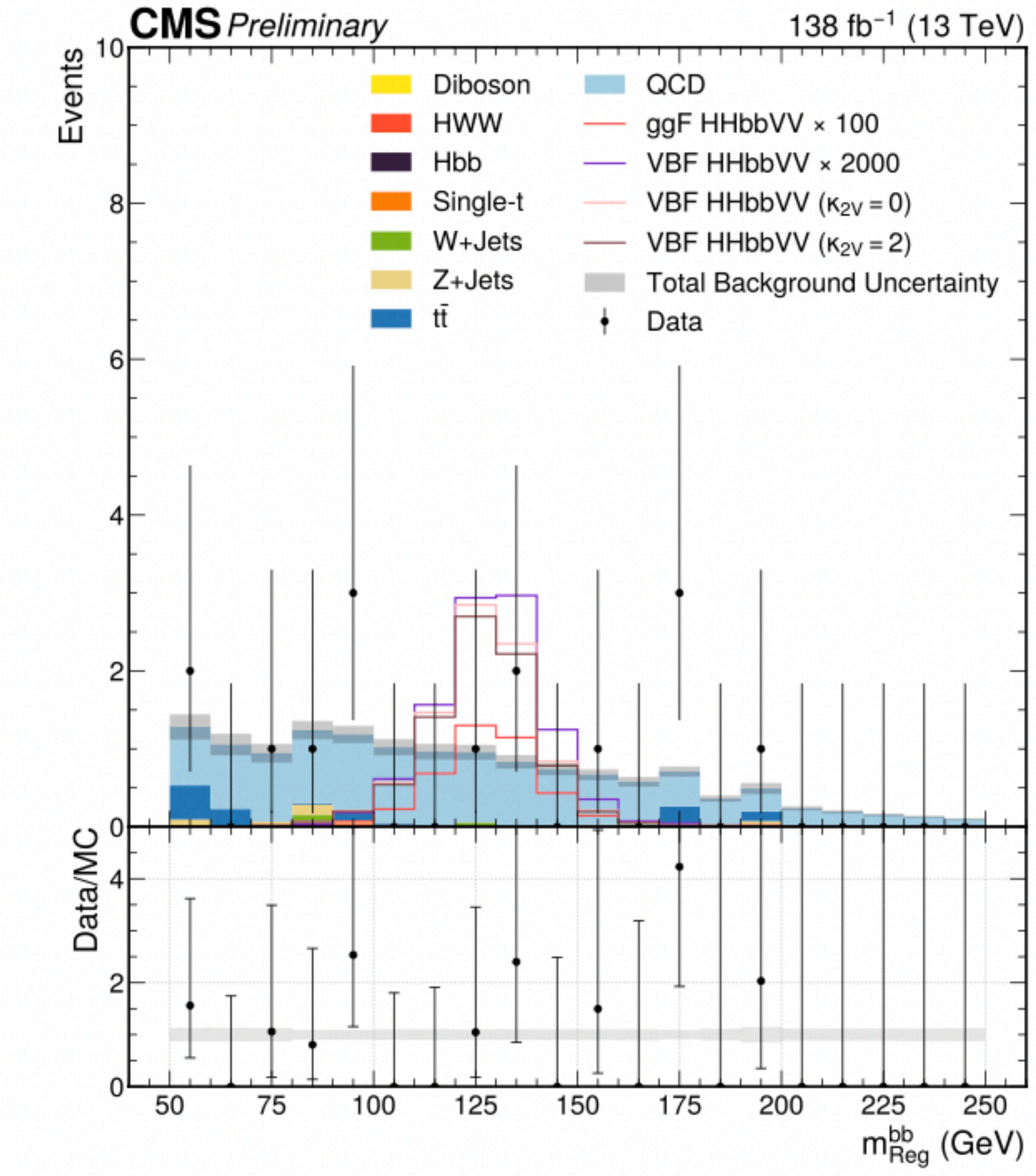
Process	Final state/ prongness	heavy flavour	# of classes
H→VV (full-hadronic)	qqqq	0c/1c/2c	3
	qqq		3
H→WW (semi-leptonic)	$e\nu qq$	0c/1c	2
	$\mu\nu qq$		2
	$\tau_e\nu qq$		2
	$\tau_\mu\nu qq$		2
	$\tau_h\nu qq$		2
H→qq		bb	1
		cc	1
		ss	1
		qq (q=u/d)	1
H→ττ	$\tau_e\tau_h$		1
	$\tau_\mu\tau_h$		1
	$\tau_h\tau_h$		1
t→bW (hadronic)	bqq	1b + 0c/1c	2
	bq		2
t→bW (leptonic)	$be\nu$	1b	1
	$b\mu\nu$		1
	$b\tau_e\nu$		1
	$b\tau_\mu\nu$		1
	$b\tau_h\nu$		1
QCD		b	1
		bb	1
		c	1
		cc	1
		others (light)	1

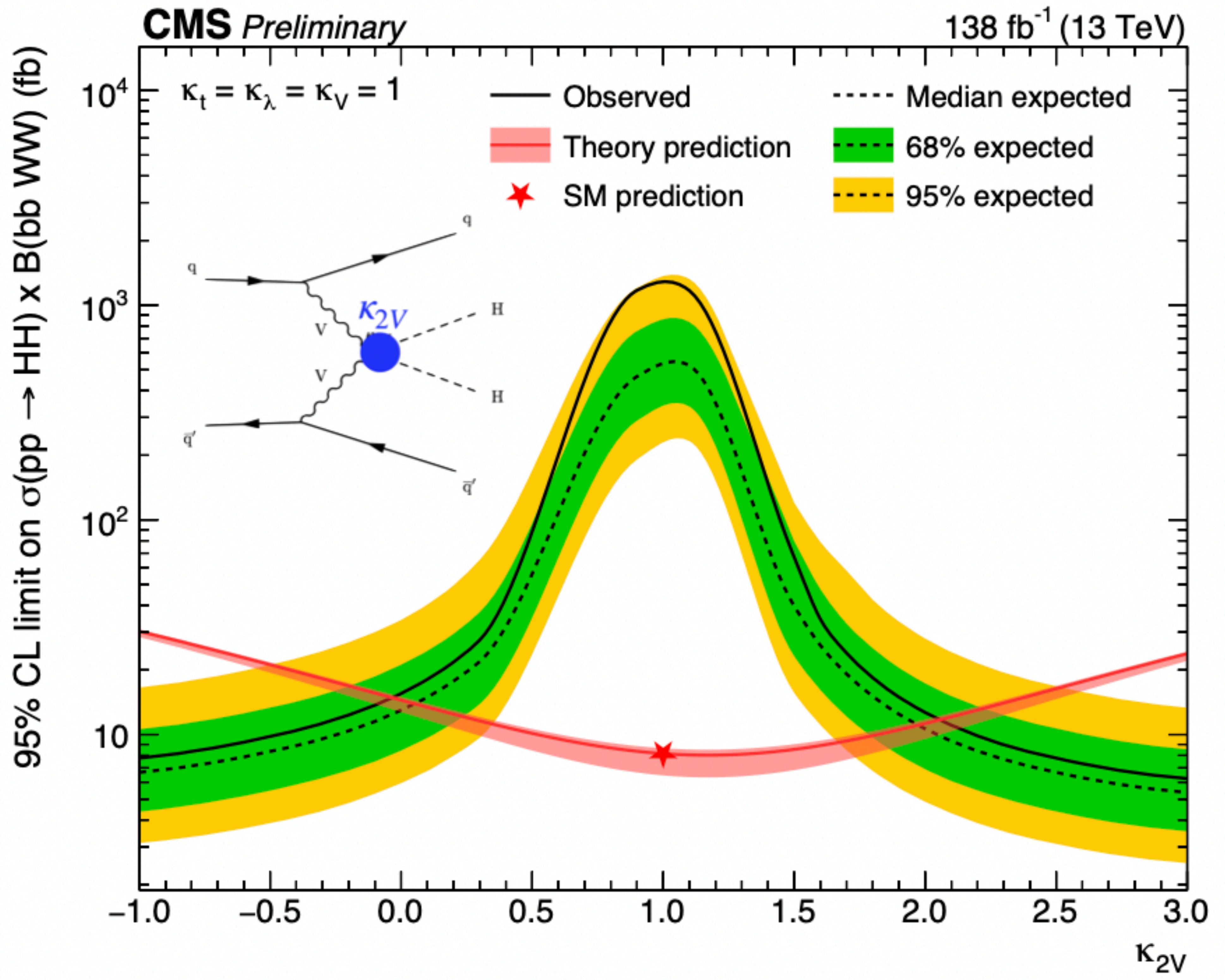


CMS-PAS-HIG-23-012

CMS non-resonant boosted HH→bbVV analysis (2024)

- Enables a new search for boosted HH → bbVV → bb4q
  - established ParticleNet mass-decorrelated tagger for H → bb jets
  - new high-performing GloParT tagger for H → VV jets
- **Provides second-best constraint on HHVV coupling K_{2V}**





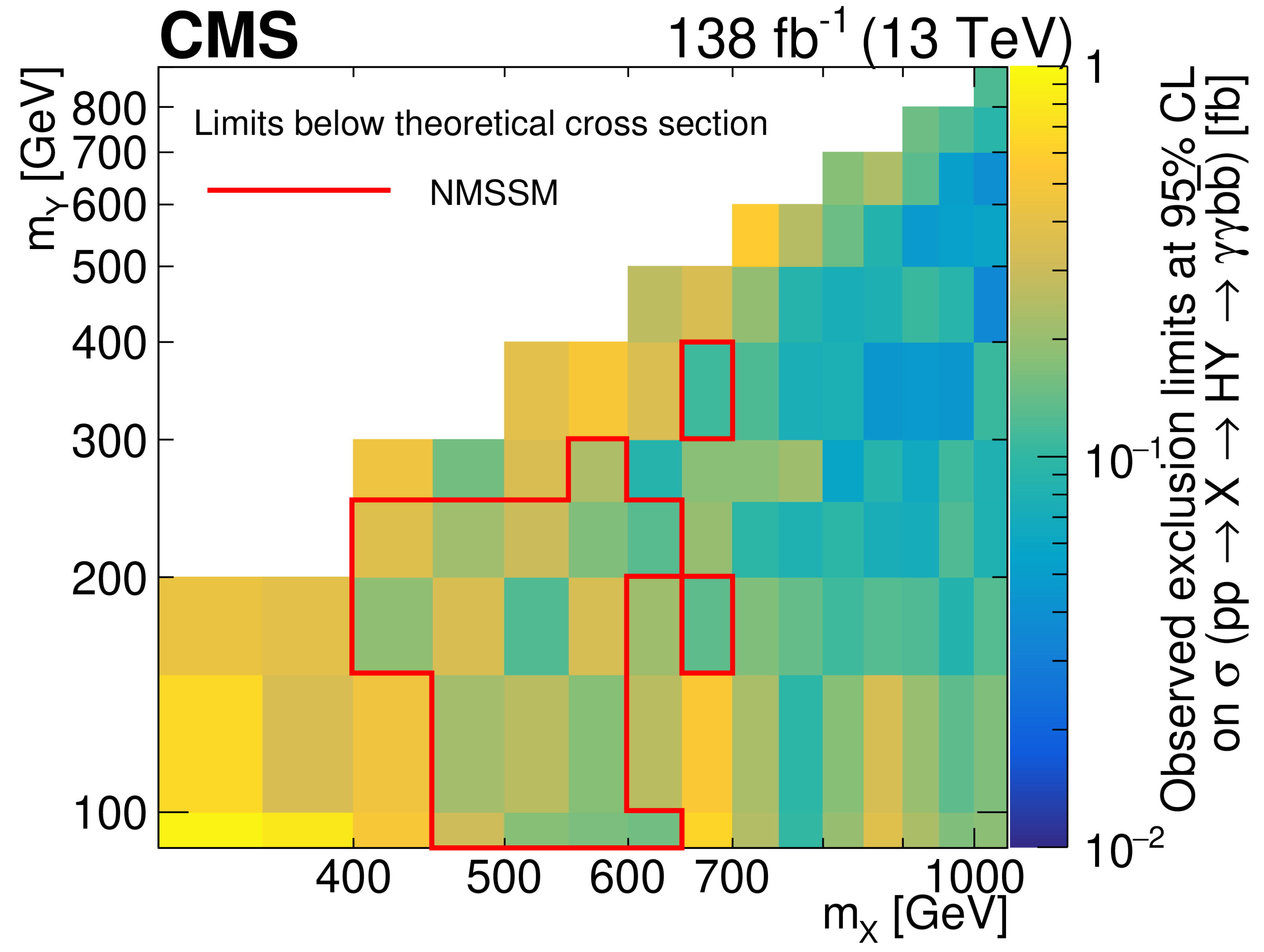
CMS-PAS-HIG-23-012

AI-based event classification in heavy resonance searches

CMS X → Y(bb)H(γγ)

	m_Y < 300 GeV	m_Y = [300–500] GeV	m_Y > 500 GeV
m_X < 500 GeV	CAT 0 = 0.63–1.0 CAT 1 = 0.33–0.63 CAT 2 = 0.17–0.33		
m_X = [500–700] GeV	CAT 0 = 0.55–1.0 CAT 1 = 0.40–0.55 CAT 2 = 0.21–0.40	CAT 0 = 0.60–1.0 CAT 1 = 0.35–0.60 CAT 2 = 0.18–0.35	
m_X > 700 GeV	CAT 0 = 0.50–1.0 CAT 1 = 0.30–0.50 CAT 2 = 0.21–0.30	CAT 0 = 0.35–1.0 CAT 1 = 0.24–0.35 CAT 2 = 0.18–0.24	CAT 0 = 0.40–1.0 CAT 1 = 0.29–0.40 CAT 2 = 0.13–0.29

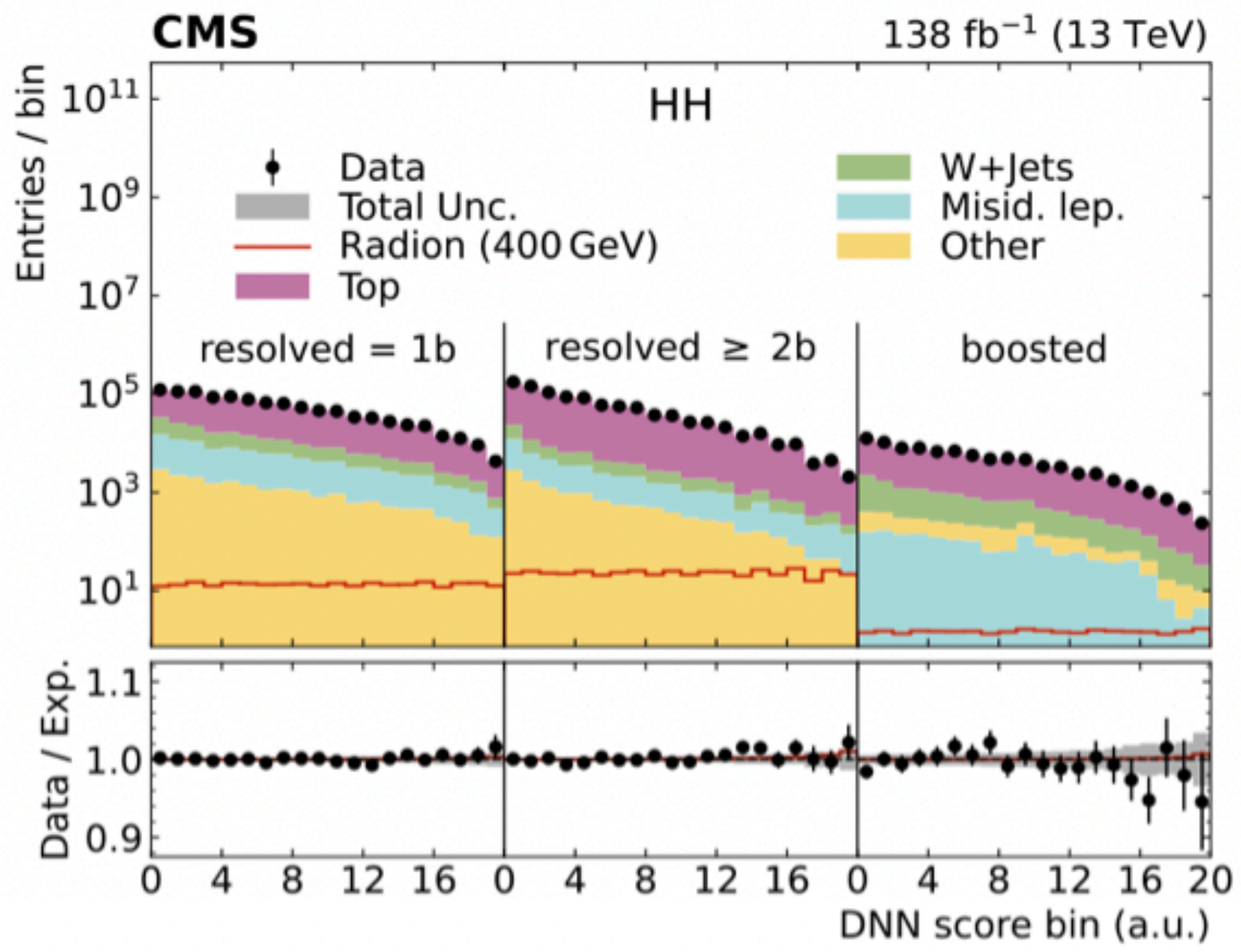
- Six exclusive kinematic regions are defined based on hypothesised values of m_X and m_Y
- In each kinematic region, a BDT with 3 output classes (2 for backgrounds and 1 for signal) is trained
  - all contained signal samples and the two background samples are used with equal weight
- In each kinematic region, 3 event categories are defined based on output of corresponding BDT
  - for each m_X hypothesis, signal is inferred from a fit in 2D distributions of $m_{\gamma\gamma}$ and m_{jj}

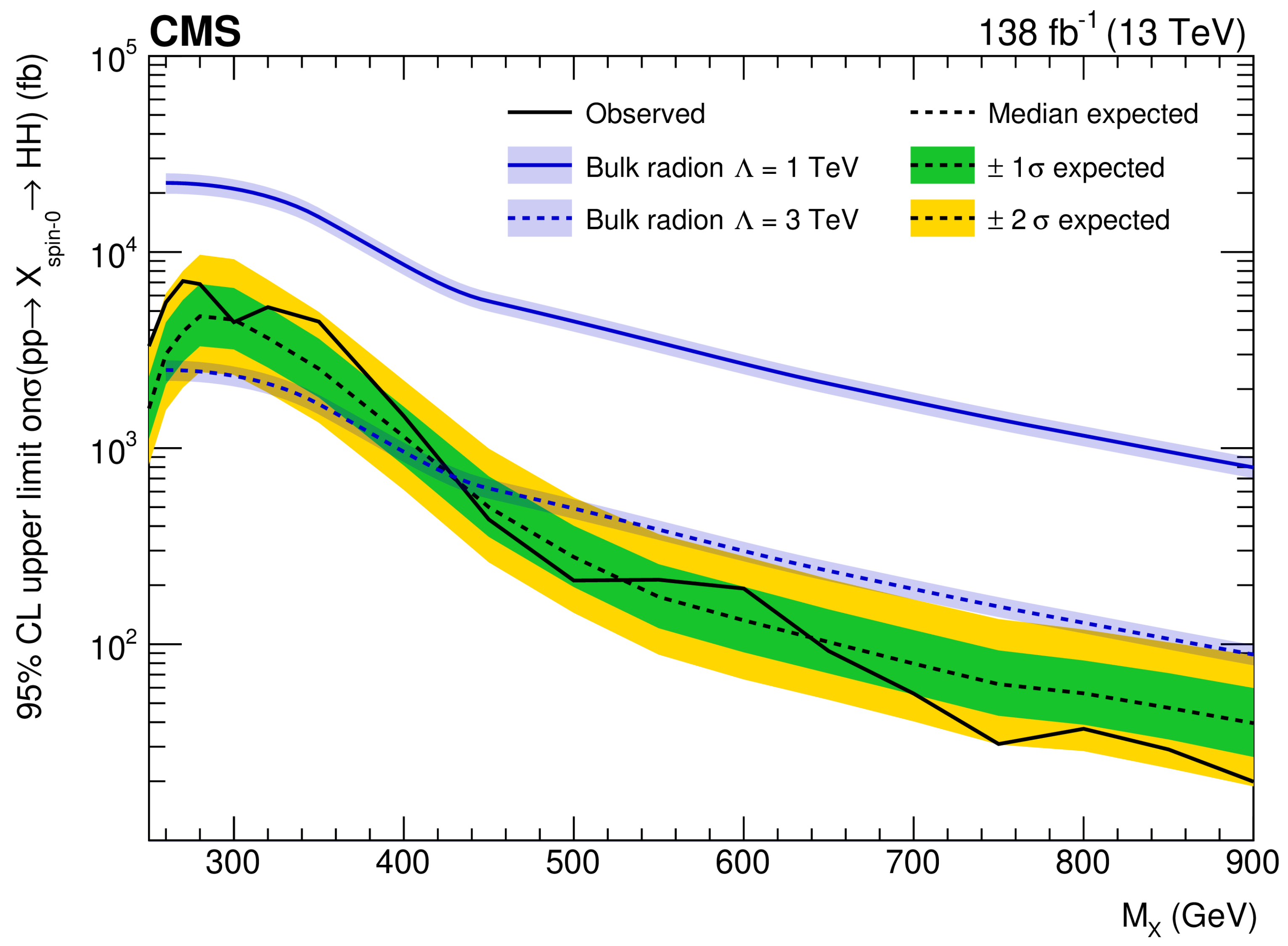


JHEP 05 (2024) 316

CMS X → H(bb)H(WW)

- DNNs feature output nodes for a number of backgrounds and one signal node
  - DNNs are trained on all signal samples; they are parameterized in nominal signal mass
- DNN architecture is complemented by a Lorentz Boost Network acting as input preprocessor
  - takes four-vectors of reconstructed particles as input and creates additional observables
- Depending on the highest scoring node, events are subdivided into signal and background categories
  - signal extraction is performed by a fit to DNN output distributions

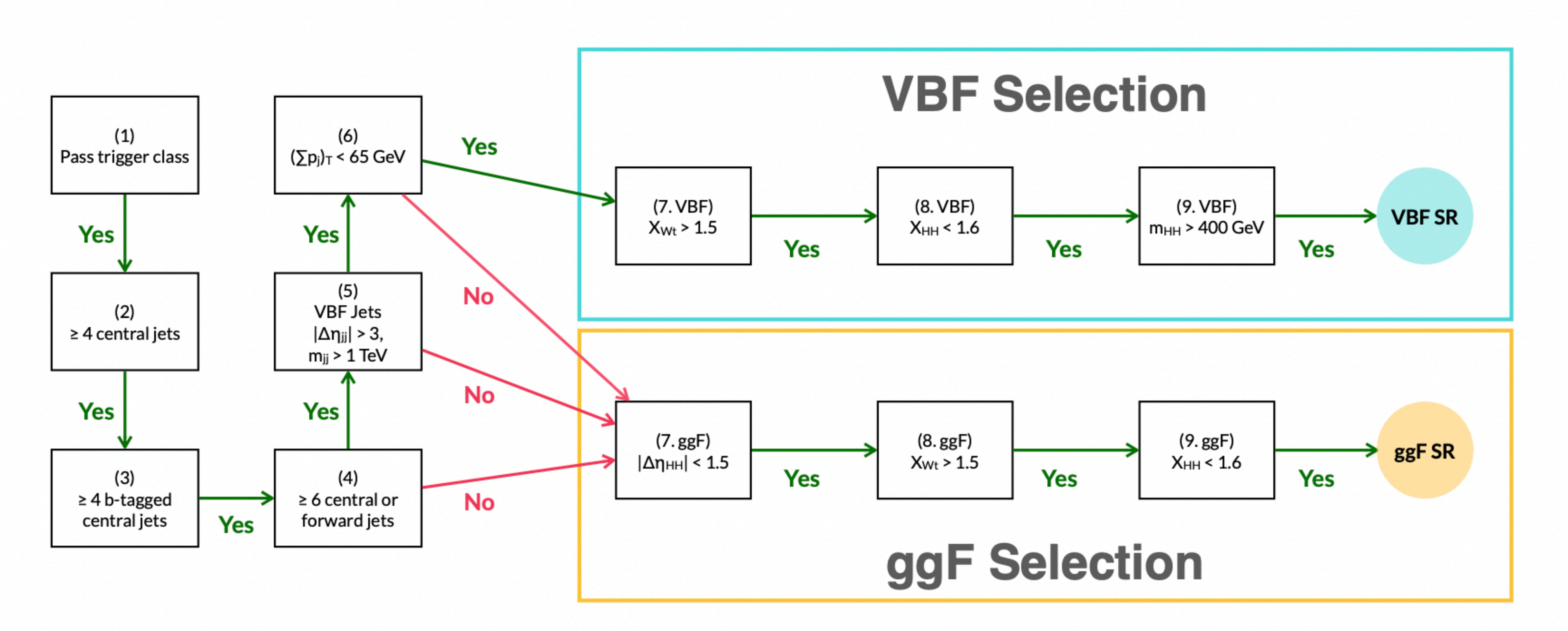




arxiv:2403.09430

ML-based background modeling

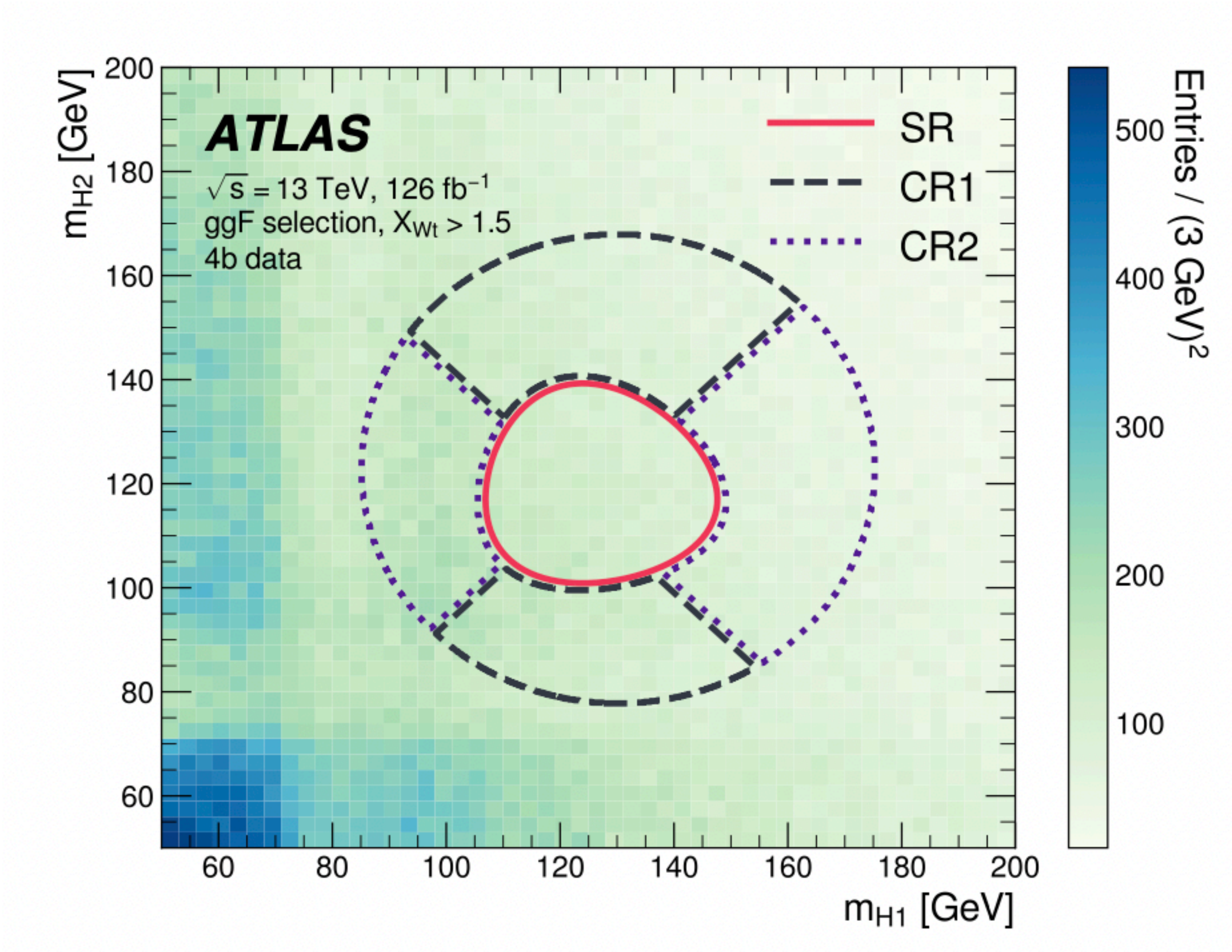
ATLAS HH→bbbb analysis (2023)

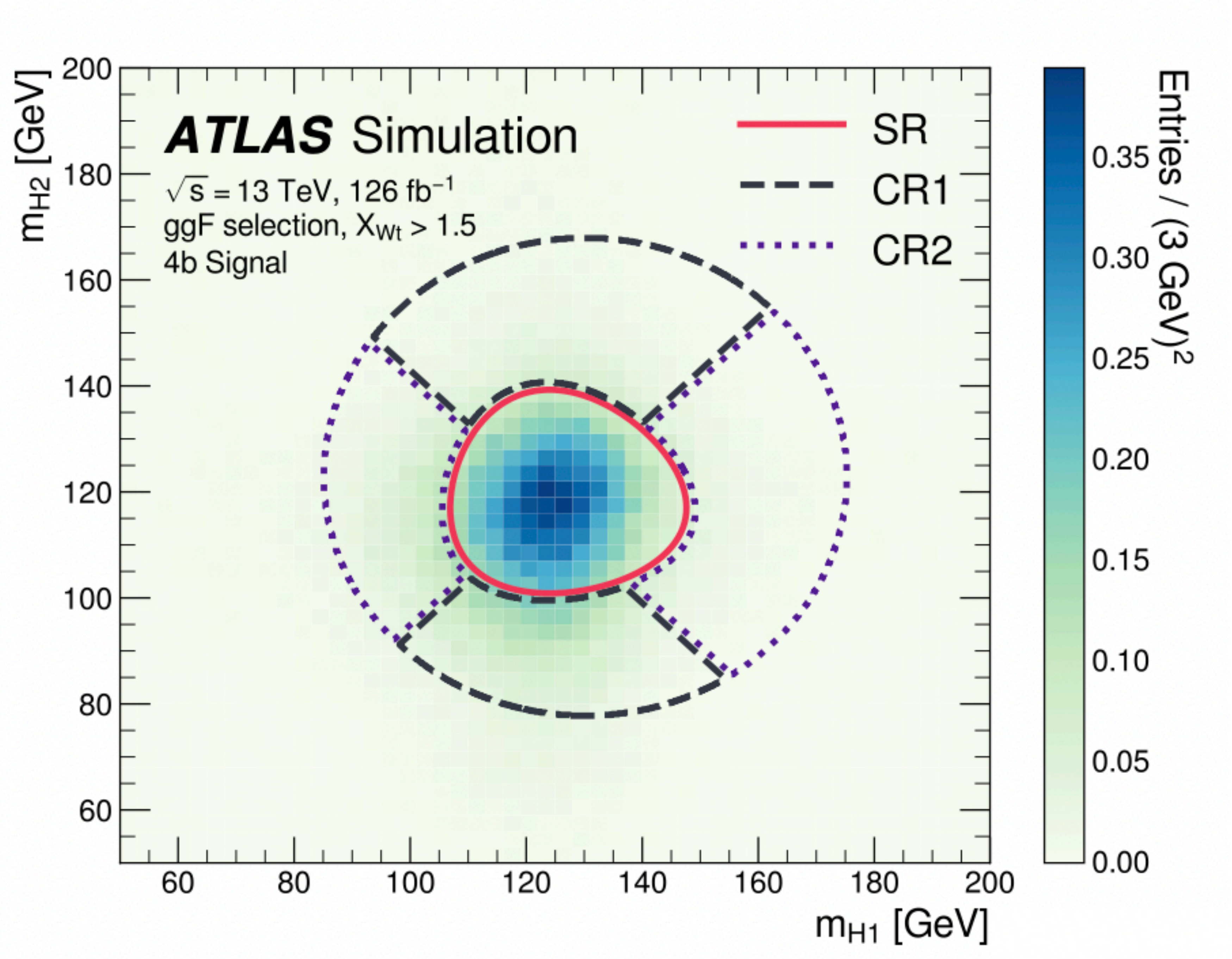


- Analysis selection: cut-based, 4b
- Background events : 90% from multi-jet and 10% from ttbar
  - modeled using a fully data-driven technique

Phys. Rev. D 108 (2023) 052003

ATLAS HH→bbbb analysis (2023)



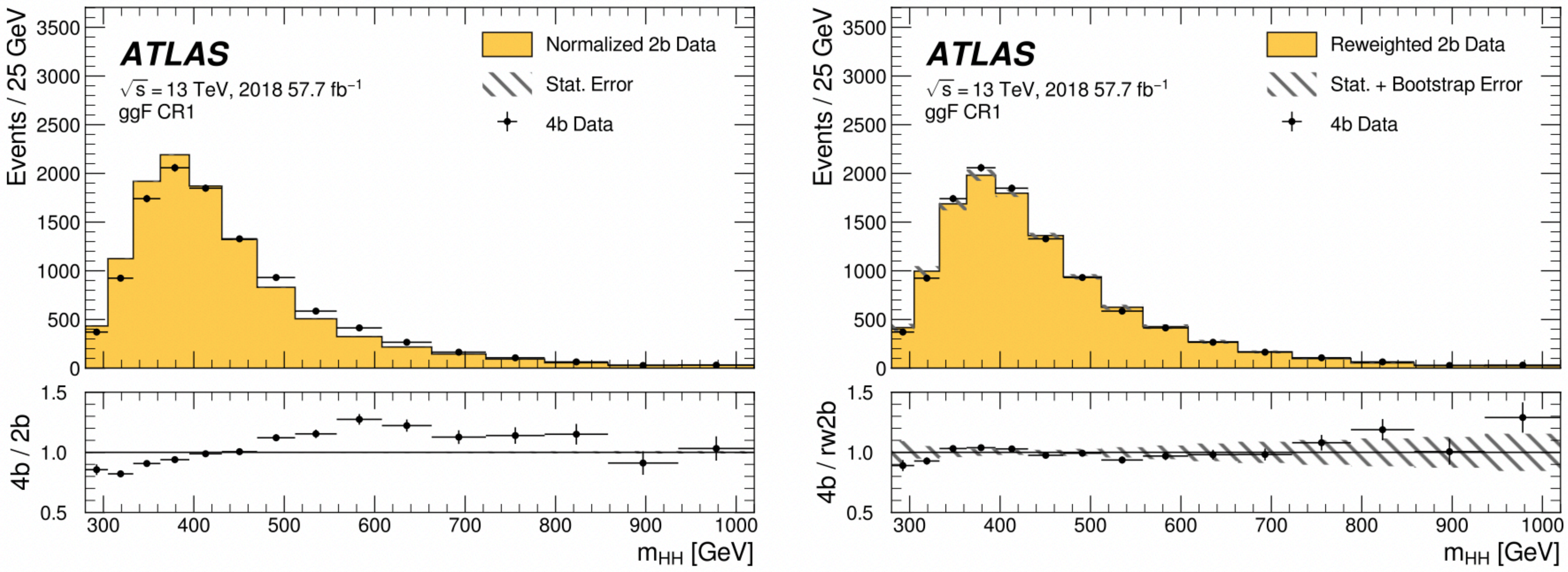


$$w(\vec{x}) = \frac{p_{4b}(\vec{x})}{p_{2b}(\vec{x})},$$

- Reweight 2b events to estimate 4b events
- Weights are derived by an artificial neural network (NN)
  - in CR1 (for nominal) and CR2 (for systematic)
  - with kinematic variables that exhibit larger differences between the $2b$ and $4b$
- To estimate systematic of varying initial conditions and limited size of training samples, construct a set of training datasets by sampling from original dataset

Phys. Rev. D 108 (2023) 052003

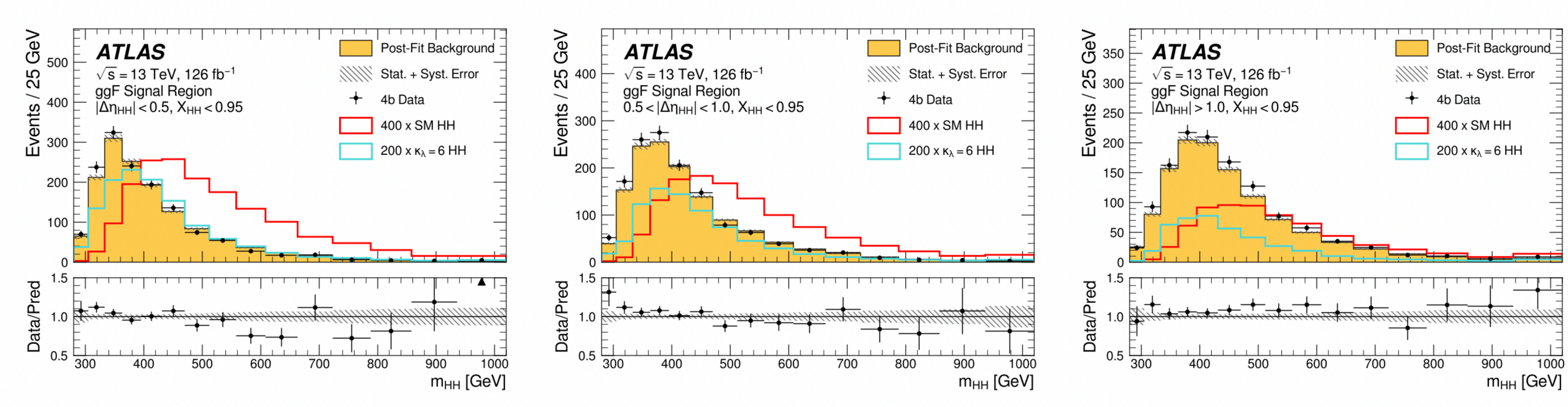
ATLAS HH→bbbb analysis (2023)



- Reweighed 2b distributions agrees 4b distributions in CR1
- Background procedure was tested with simulation samples
- Also tested in several control data samples
  - 2*b* and 4*b* events with $|\Delta\eta_{HH}| > 1.5$
  - 2*b* and 4*b* events with shifted center of SR
  - events with exactly 3 *b*-tagged jets plus 1 central jet failing *b*-tagging requirement

Phys. Rev. D 108 (2023) 052003

ATLAS HH→bbbb analysis (2023)



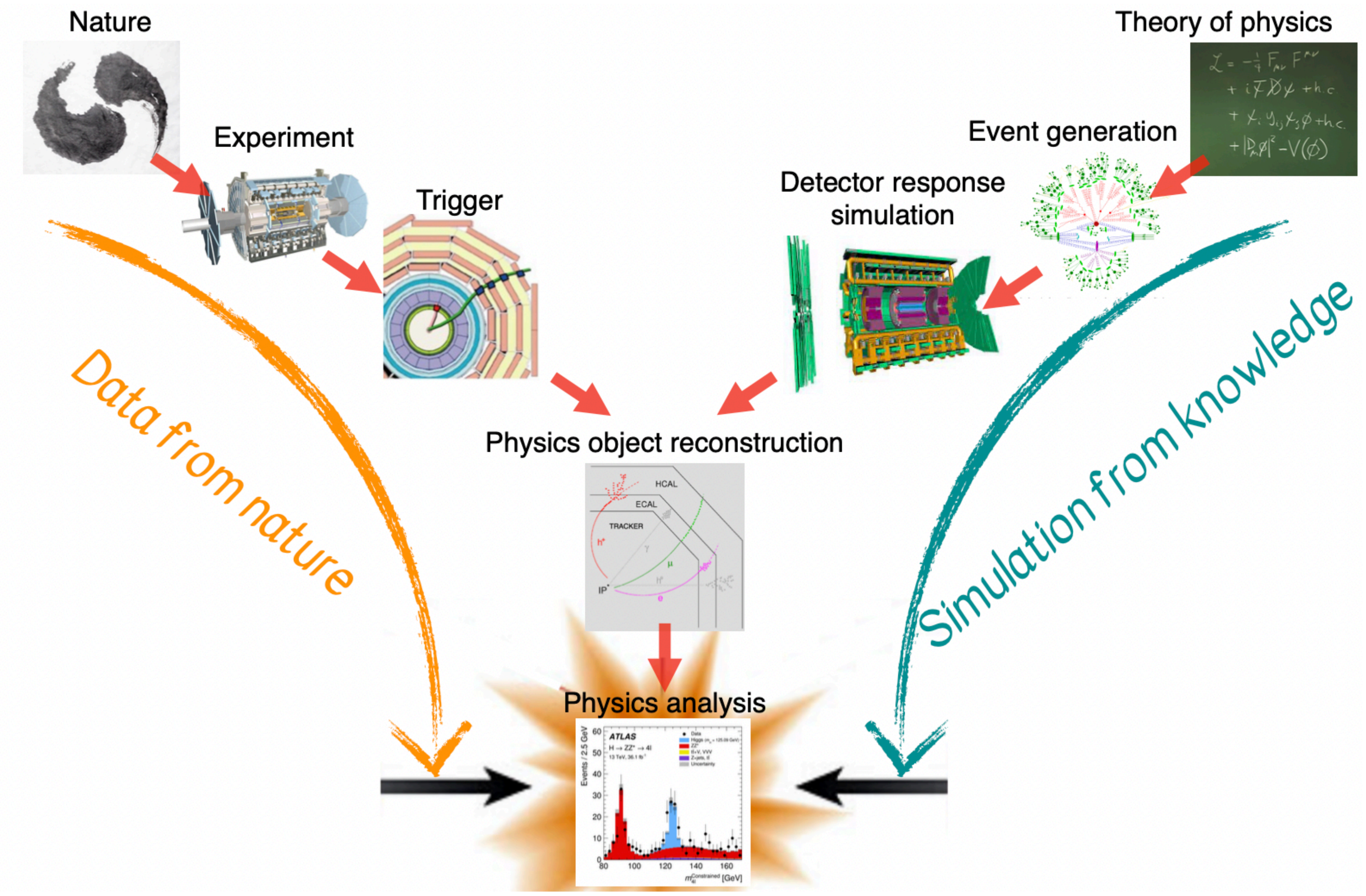
- “The sensitivity of the analyses is improved relative to previous iterations by using more sophisticated background modeling techniques…” :-)
- No evidence of signal is found :-(

Source of Uncertainty	$\Delta\mu/\mu$
Theory uncertainties	
Theory uncertainty in signal cross-section	−9.0%
All other theory uncertainties	−1.4%
Background modeling uncertainties	
Bootstrap uncertainty	−7.1%
CR to SR extrapolation uncertainty	−7.5%
$3b1$f nonclosure uncertainty	−2.0%

Phys. Rev. D 108 (2023) 052003

Summary

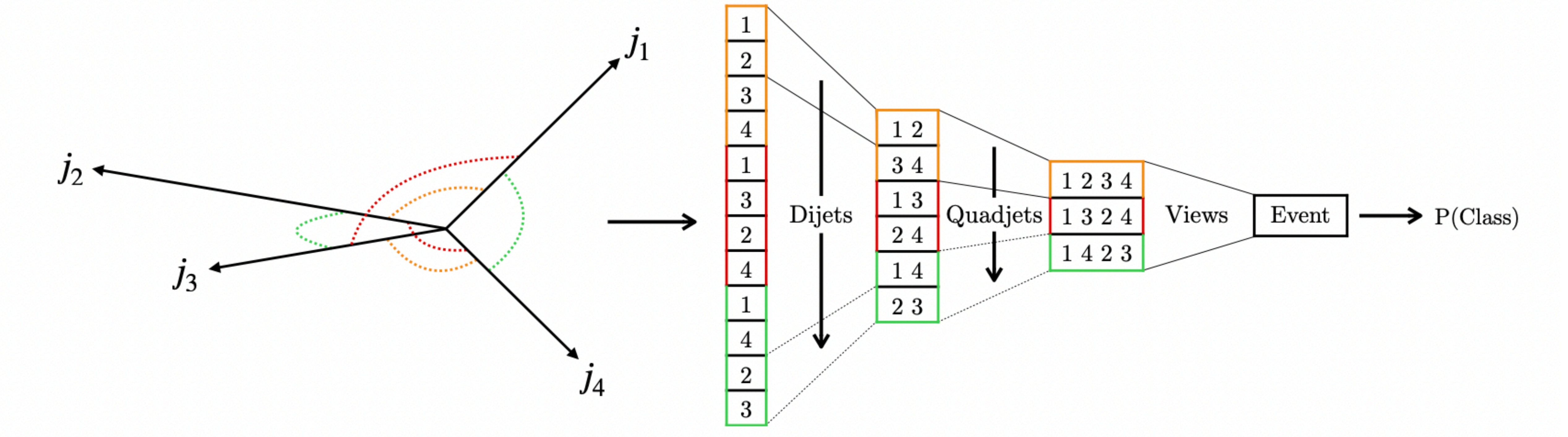
- **Machine Learning greatly enhances our ability of identifying signal from background: important for discovery of HH**
- Lots of recent progress at ATLAS and CMS:
  - deep learning particle/event reconstruction
  - ML-based background modeling
  - etc.
- And there are much more to come!



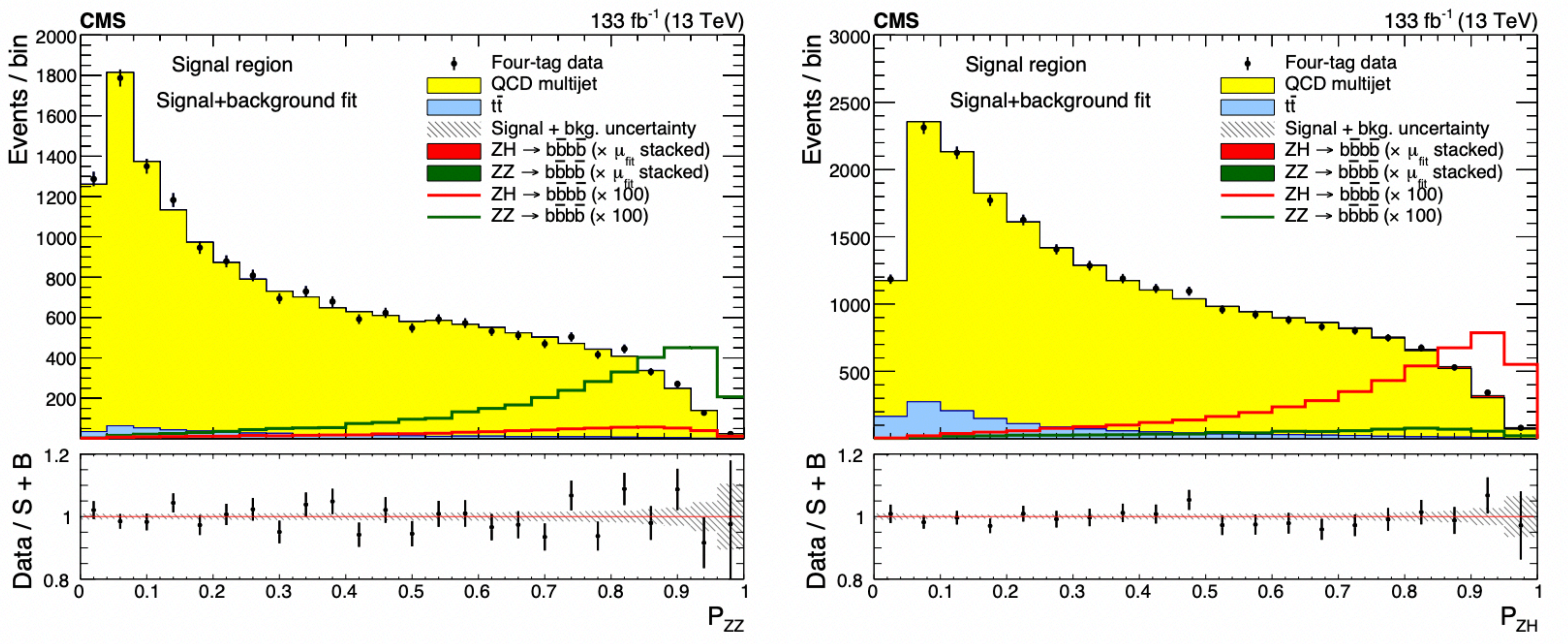
Thanks!

ZZ/ZH → 4b

- Search for ZZ and ZH production in 4b final state
- Benefits from a multiclass multivariate classifier, which uses convolutions to solve combinatoric jet pairing problem, and has been designed with an architecture customized to 4b final state



- Observed (expected) upper limits on ZZ → 4b and ZH → 4b production cross sections correspond to 3.8 (3.8) and 5.0 (2.9) times SM prediction, respectively
- Analysis techniques directly applicable to the HH → 4b analysis



arxiv:2403.20241